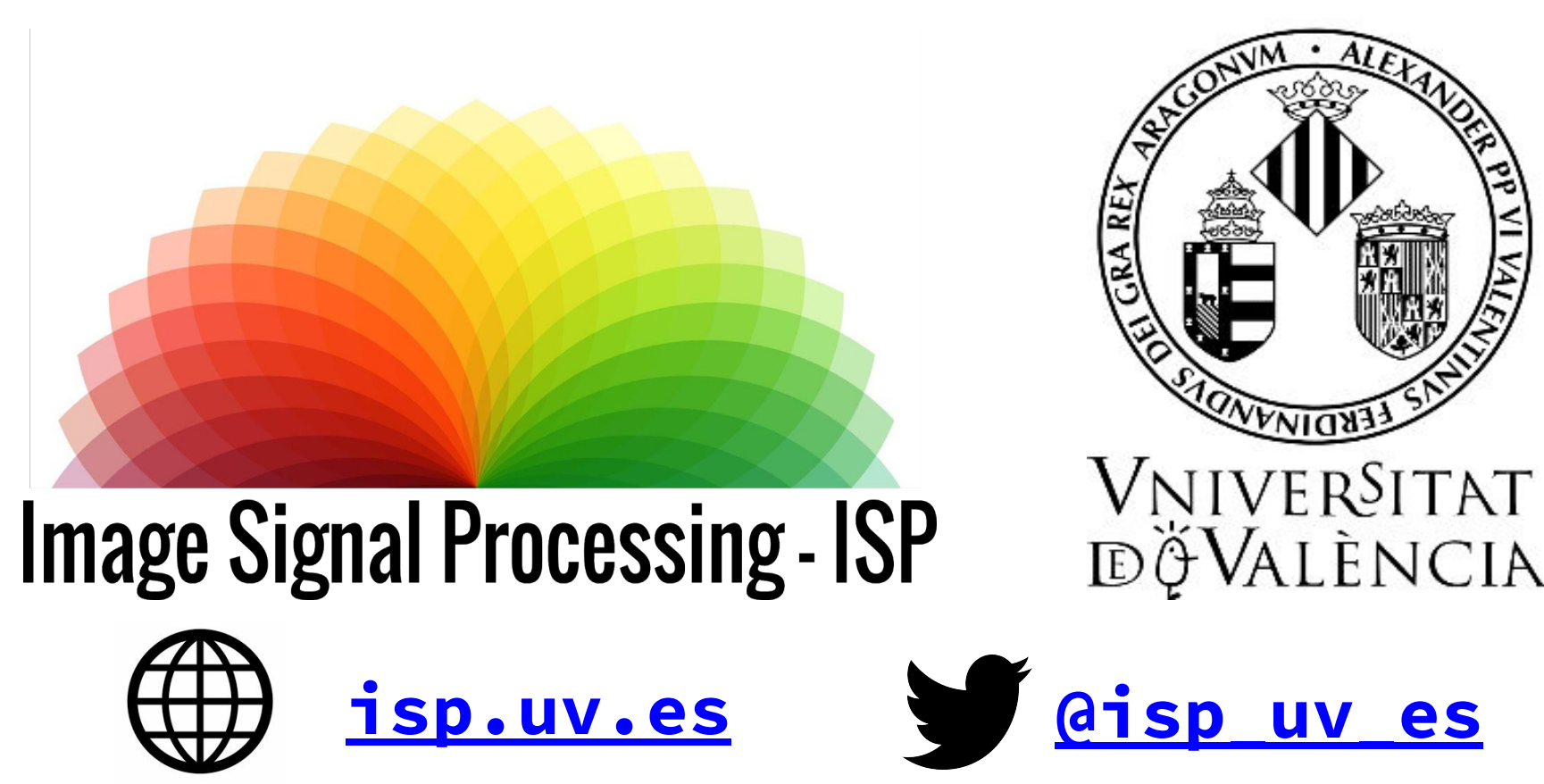


Gaussianization for Multivariate, High-dimensional Earth Observation data Analysis

J. Emmanuel Johnson, Maria Piles, Valero Laparra, and Gustau Camps-Valls

juan.johnson@uv.es
[@iejjohnson](#)
University of Valencia, Image and Signal Processing Laboratory Paterna, Spain

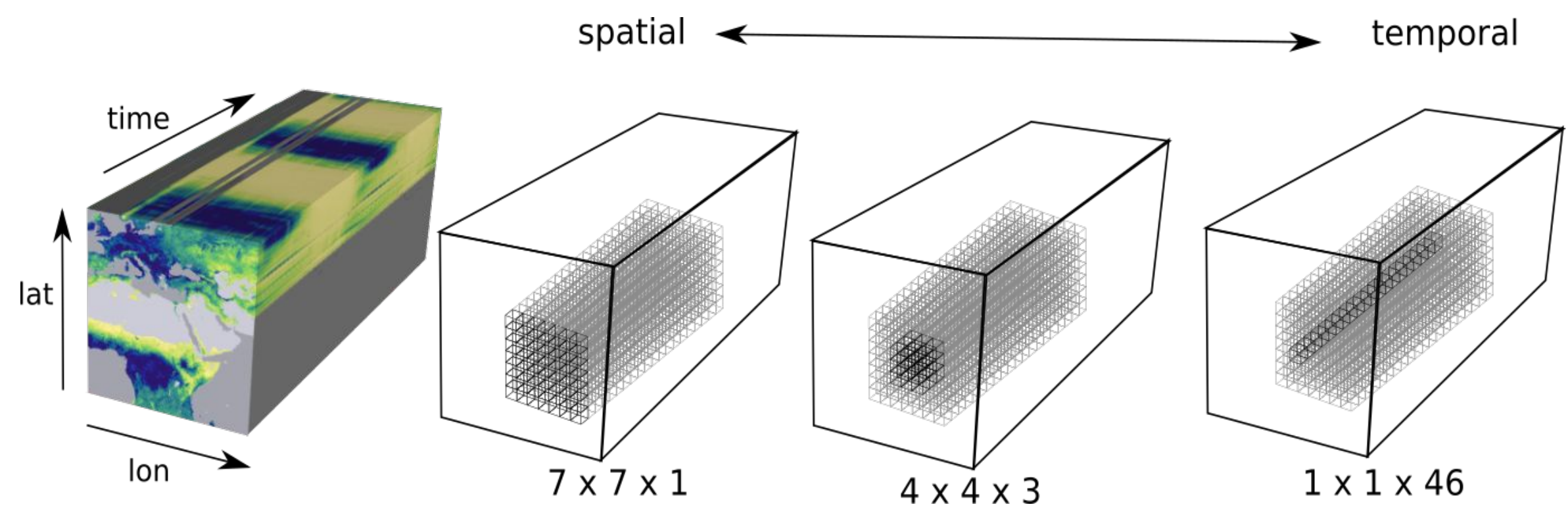


How do we represent spatial-temporal data?

Information Theory (IT) provides **ideal measures** for choosing an appropriate representation. IT measures allow us to:

- Quantify Uncertainty
- Non-Linear Relationships
- Naturally Multivariate
- Universal Units

However, it's not widely adopted due to the problem of **density estimation** for high dimensional, multivariate data.

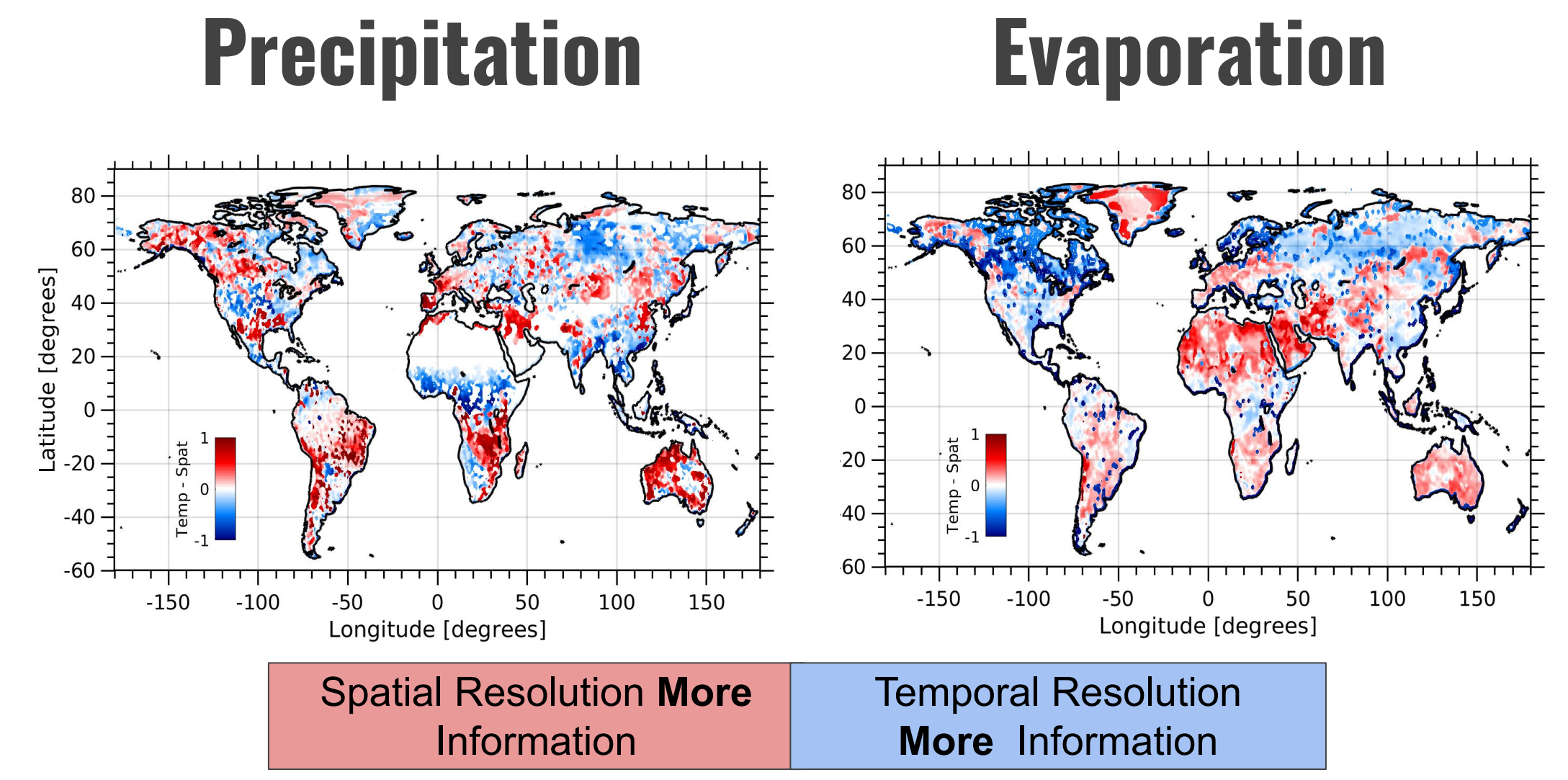


Density Estimation

Although Gaussianization is a simple formulation, it is a **scalable** and **robust** density estimator [1].

We demonstrate its capabilities for:

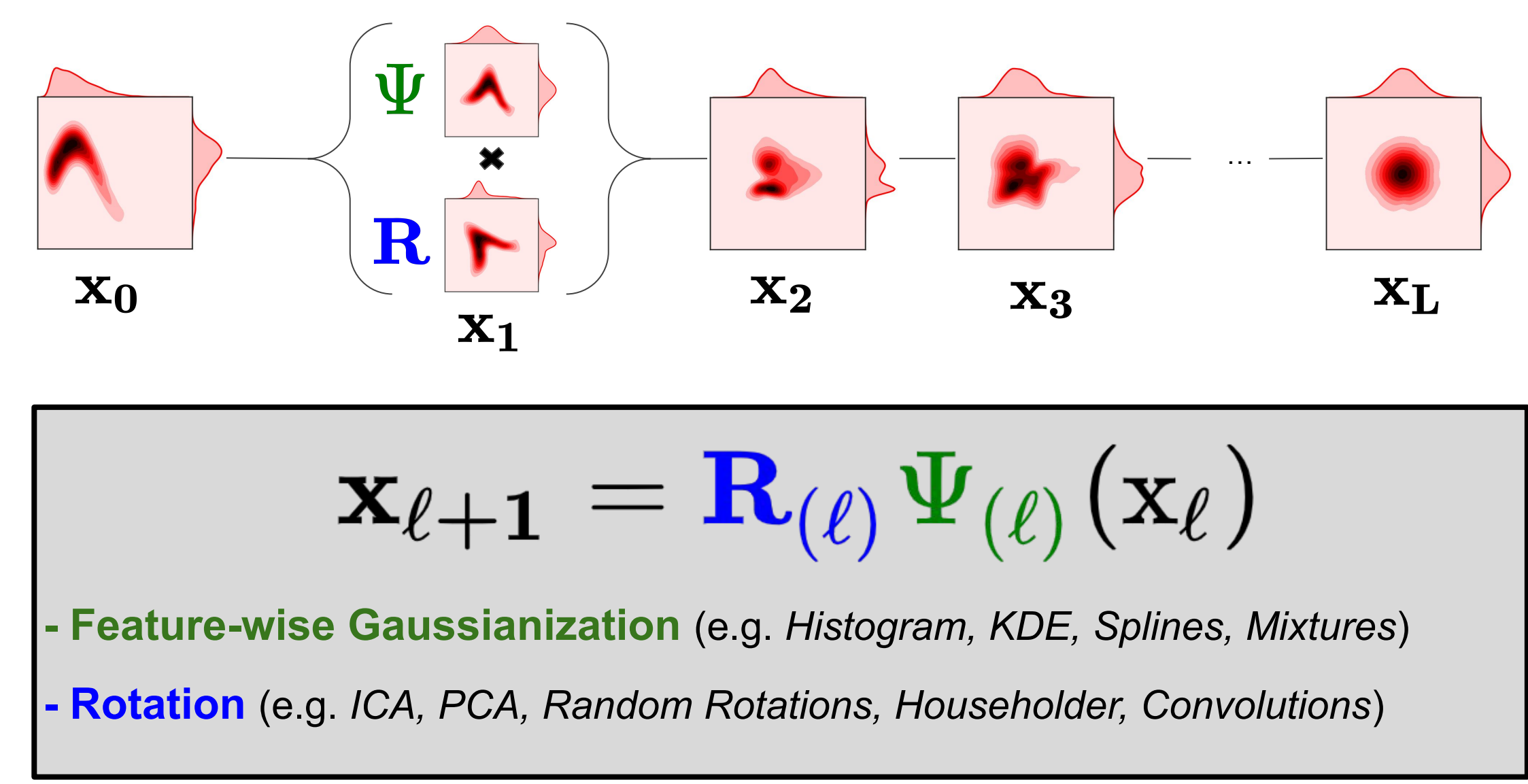
- RADAR data
- Hyperspectral Image data
- Drought Variables
- Earth System Data (shown here)



Gaussianization (aka Normalizing Flows)

Normalizing Flows are fully parametric, **generative models** composed of a sequence of **invertible transformations**. Then, by using the change of variables formula, one can calculate **exact likelihoods** as well as **generate samples**.

Gaussianization is a particular (and very simple) formulation of this family which features a parametric and iterative form.



Information Theory Metrics

Gaussianization has deep connections to **information theory** [1, 2]. This allows for **fast** and **robust** estimation of IT metrics for high-dimensional, multivariate data.

We show how experiments for:

- Entropy (shown here)
- Total Correlation
- Multivariate Mutual Information
- Kullback-Leibler Divergence

